APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/748,453	12/26/2000	Mei-Yuh Hwang	M61.12-0339	3442

7590 12/09/2003	EXAMINER
WESTMAN, CHAMPLIN & KELLY, P.A.	NOLAN, DANIEL A
Suite 1600 - International Center	

ART UNIT	PAPER NUMBER
2654	12

900 Second Avenue South
Minneapolis, MN 55402-3319

DATE MAILED: 12/09/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on _30 September 2003 & 17 November 2003_ .

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) _1-21_ is/are pending in the application.

     4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) _1-21_ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☒ The proposed drawing correction filed on _09 June 2003_ is: a)☒ approved b)☐ disapproved by the Examiner.

     If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

     a)☐ All b)☐ Some * c)☐ None of:

         1.☐ Certified copies of the priority documents have been received.

         2.☐ Certified copies of the priority documents have been received in Application No. _____ .

         3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

     * See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

     a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _2,3_ .

4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____ .

DETAILED ACTION

1.   The text of those sections of Title 35, U.S. Code not included in this action can

be found in a prior Office action.

Continued Examination Under 37 CFR 1.114

2.   A request for continued examination under 37 CFR 1.114, including the fee set

forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this

application is eligible for continued examination under 37 CFR 1.114, and the fee set

forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action

has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 17

November 2003 has been entered.

Response to Amendment

3.   The response of 30 September 2003 was entered to the effect that the claims

were changed as indicated and examined on the merits.

Response to Arguments

4.   Applicant's arguments filed 30 September have been fully considered but they are not persuasive and the rejection of claim 19 is maintained.

In response to applicant's argument that Schultze is not applicable because language is unknown, the fact that applicant has recognized another advantage which would flow naturally from following the suggestion of the prior art cannot be the basis for patentability when the differences would otherwise be obvious. See *Ex parte Obiaya*, 227 USPQ 58, 60 (Bd. Pat. App. & Inter. 1985).

In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., *rules for identifying trigrams*) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

With respect to that issue, without further specification in the application, *syllable-like units* is understood to be as established by the preamble of Gupta *et al* establishing that during processing (column 3 lines 53-67), *symbols that compose a transcription may be phonemes, allophones, triphones, syllables and demi-syllables* (column 3 lines 59-60), making such elements correspond to the syllable-like units being processed by claim 19.

Claim Rejections - 35 USC § 103

5.   This application currently names joint inventors. In considering patentability of

the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of

the various claims was commonly owned at the time any inventions cover ed therein

were made absent any evidence to the contrary. Applicant is advised of the obligation

under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was

not commonly owned at the time a later invention was made in order for the examiner to

consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g)

prior art under 35 U.S.C. 103(a).

Smith *et al* & Häb-Umbach *et al*

6.   Claims 1–4 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Smith *et al* (U.S. Patent 6,408,271 B) in view of Häb-Umbach *et al* (U.S. Patent

5,873,061 A).

7.   Regarding claim 1, the features employed by Smith *et al* in *generating phrasal*

transcriptions for speech recognition dictionaries by permutating word transcriptions for

each vocabulary item in an orthographic phrase read on the features of the *method for*

adding an acoustic description of a word to a speech recognition lexicon of the

immediate application as follows:

- Smith *et al* (column 6 lines 15-20) reads on the feature of *converting the text of the word into at least one orthographically derived acoustic description of the word;*

- Smith *et al* (column 6 lines 42-46) reads on the feature of *generating a score for an orthographically derived acoustic description based in part on a comparison between the orthographically derived acoustic description and a speech signal representing a user's pronunciation of the word;*

- Smith *et al* (with *generating* steps 202 & 302 in figures 2 & 3) reads on the feature of identifying a speech-based acoustic description 804 in figure 8) *representing the user's pronunciation of the word to produce a* speech-based *acoustic description of the word and a score for the* speech-based *acoustic description; and*

- Smith *et al* (column 12 lines 26-37) reads on the feature of *selecting one of the orthographically derived acoustic description and the* speech-based *acoustic description as the acoustic description of the word based on the score for the orthographically* (column 12 lines 30-31) *derived acoustic description and the score for the* speech-based *acoustic description* (column 12 lines 35-37).

Smith *et al* does not stipulate that the *acoustic description* be either speech-based or that it is *not associated with the text of the word*. The invention *for constructing a model of a new word for addition to a word model database of a speech recognition system* of Häb-Umbach *et al* (figures 1 & 2) reads on the limitation that the *acoustic description is speech-based* (column 3 lines 14-17) and the step of identifying a speech-based acoustic description *from the speech signal representing the user's pronunciation of the word, wherein the speech-based acoustic description is not*

associated with the text of the word (column 4 lines 35-42). It would have been obvious

to a person of ordinary skill in the art of speech signal processing at the time of the

invention to apply the method/teachings of Häb-Umbach *et al* to the device/method of

Smith *et al* so as to make recognition not be dependent on transcription by having any

subsequent transcription be a separate independent operation apart from of recognition.

8.   Regarding claim 2, the claim is set forth with the same limits as claim 1.

Smith *et al* (column 13 lines 53-56) reads on the feature of *generating an acoustic*

model score.

9.   Regarding claim 3, the claim is set forth with the same limits as claim 2.

Smith *et al* (column 13 lines 57-60) reads on the feature of *generating an acoustic*

model score for at least one speech-based *acoustic description and using the score as*

at least part of the score for the speech-based *acoustic description.*

10.   Regarding claim 4, the claim is set forth with the same limits as claim 3.

Smith *et al* (802 in figure 8) reads on the feature of *using the same acoustic model*

(specified by *"a speech model set"*, column 13 line 52) *to generate both acoustic model*

scores (lines 46-56).

Smith *et al*, Häb-Umbach *et al* & Bahl *et al*[']426

11.   Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over Smith *et al* in view of Häb-Umbach *et al* and further in view of Bahl *et al*[']426 (U.S. Patent 5,875,426).

12.   Regarding claim 5, the claim is set forth with the same limits as claim 3. Neither Smith *et al* nor Häb-Umbach *et al* teach *generating a language model score.* The Bahl *et al*[']426 method/system for *recognizing speech having word liaisons by adding a phoneme to reference word models* (column 3 lines 55-60) reads on the feature of *generating a language model score for the at least one* speech-based *acoustic description and* (lines 58-59) *using the language model score as part of the score for the at least one* speech-based *acoustic description.*

It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Bahl *et al*[']426 to the device/method of Smith *et al* & Häb-Umbach *et al* so as to consider context among the bases of making an acoustic decision.

Smith *et al*, Bahl *et al*[']426 & Bahl *et al*[']921

13.   Claims 6–8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Smith *et al* in view of Häb-Umbach *et al* and further in view of Bahl *et al*[']426 and further in view of Bahl *et al*[']921 (U.S. Patent 6,377,921).

14.   Regarding claim 6, the claim is set forth with the same limits as claim 5.

Neither Smith et al nor Häb-Umbach et al teach generating a language model score.

The Bahl et al[921] method/system for identifying mismatches between assumed and

actual pronunciations of words (column 2 lines 30-38) reads on the feature of

generating an acoustic model score and a language model score (lines 39-40 from 18B

in figure 1) for a sequence of syllable-like units (column 5 lines 47-48) and (with lines

55-61) the further feature that the speech-based acoustic description is derived from the

sequence of syllable-like units.

It would have been obvious to a person of ordinary skill in the art of speech

signal processing at the time of the invention to apply the method/teachings of Bahl et

al[921] to the device/method of Smith et al & Häb-Umbach et al to increase precision and

avoid prosodic differences by addressing the lower cohesive elements of speech.

15.   Regarding claim 7, the claim is set forth with the same limits as claim 6.

Neither Smith et al nor Häb-Umbach et al teach generating a language model score.

Bahl et al[921] (with the "phones" of column 6 line 13) reads on the feature of dividing the

sequence of syllable-like units into a sequence of phonemes, which would have made it

obvious to a person of ordinary skill in the art of speech signal processing at the time of

the invention to apply the method/teachings of Bahl et al[921] to the device/method of

Smith et al so as to not to over look minor utterances by considering each potential

word segment separately.

16.   Regarding claim 8, the claim is set forth with the same limits as claim 6.

Neither Smith et al nor Häb-Umbach et al teach generating a language model score.

Bahl et al[426] (column 3 lines 51-53) reads on the feature of generating a language

model score based on a trigram language model for syllable-like units, which would

have made it obvious to a person of ordinary skill in the art of speech signal processing

at the time of the invention to apply the method/teachings of Bahl et al[426] to the

device/method of Smith et al & Häb-Umbach et al so as to more quickly isolate

candidates from combinations of segments.

Smith et al, Häb-Umbach et al, Bahl et al[426] & Contolini et al

17.   Claims 9–11 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Smith et al in view of Häb-Umbach et al and further in view of Bahl et al[426] and further in

view of Contolini et al (U.S. Patent 6,233,553 B1).

18.   Regarding claim 9, the claim is set forth with the same limits as claim 6.

Neither Smith et al nor Häb-Umbach et al teach generating a language model score.

Contolini et al, in the method and system for automatically determining phonetic

transcriptions associated with spelled words, reads on the feature of generating

acoustic model (of claim 4 limiting by claim 1) scores for each of a sequence of

phonemes (column 7 line 6) that form the sequence of syllable-like units (column 6 line

56).

It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Contolini *et al* to the device/method of Smith *et al* & Häb-Umbach *et al* so as to be able to relate the results of the recognition that might require correction to those elements that would be familiar to the speaker.

19.   Regarding claim 10, the claim is set forth with the same limits as claim 1. Neither Smith *et al* nor Häb-Umbach *et al* specify the product reaching a state accessible for human intervention. Contolini *et al* does so permit such adjustments, with (figure 2) reading on the feature of *displaying a user interface comprising an edit box* (item 35) *in which a user may enter the text of the word* (as according to the 1st lines of the Abstract) *and a list box* (item 34) *that displays words for which an acoustic description has been previously added to the speech recognition lexicon.*

It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Contolini *et al* to the device/method of Smith *et al* & Häb-Umbach *et al* so as to permit refinements that recognize exceptions to the rules used to set up the vocabulary.

20.   Regarding claim 11, the claim is set forth with the same limits as claim 10. Neither Smith *et al* nor Häb-Umbach *et al* specify the product reaching a state accessible for human intervention.

- Contolini *et al* (figure 2 & column 4 lines 17-25) reads on the features of *receiving an indication that a user has selected a word in the list box* (line 22);

- Contolini *et al* (column 5 lines 55-56) reads on the features of *retrieving the added acoustic description of the word from the speech recognition lexicon and converting the retrieved acoustic description into an audible signal.*

- It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Contolini *et al* to the device/method of Smith *et al* & Häb-Umbach *et al* so as to audibly confirm the validity of the revision.

Gupta *et al* & Häb-Umbach *et al*

21.   Claim 12–17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gupta *et al* (U.S. Patent 6,243,680 B1) in view of Häb-Umbach *et al*.

22.   Regarding claim 12, the *apparatus* of Gupta *et al for obtaining a transcription of phrases through text and spoken utterances* relates to the features for *a computer-readable medium* of the immediate application as follows:

- Gupta *et al* (column 1 lines 56-57) reads on the feature of *receiving text of a word* (column 3 lines 27-28) *for which a phonetic description is to be added to a speech recognition lexicon* (line 56) and on the feature of *receiving a representation of a speech signal produced by a person pronouncing the word* (line 57);

- Gupta *et al* (412 → 400 in figure 4) reads on the feature of *converting the text of the word into a text-based phonetic description of the word;*

- Gupta *et al* (402 → 404 in figure 4) reads on the feature of *generating a speech-based phonetic description of the word from the representation of the speech signal; and*

- Gupta *et al* (406 in figure 4) reads on the feature of *selecting a phonetic description of the word to add to the speech recognition lexicon by selecting between the text-based phonetic description and the speech-based phonetic description based in part on the correspondence between each phonetic description and the representation of the speech signal.*

Gupta *et al* does not stipulate that the *acoustic description* be either speech-based or that it is *not associated with the text of the word*. The invention *for constructing a model of a new word for addition to a word model database of a speech recognition system* of Häb-Umbach *et al* (figures 1 & 2) reads on the limitation that the *acoustic description is speech-based* (column 3 lines 14-17) and the step of identifying a speech-based acoustic description *from the speech signal representing the user's pronunciation of the word, wherein the speech-based acoustic description is not associated with the text of the word* (column 4 lines 35-42). It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Häb-Umbach *et al* to the device/method of Gupta *et al* so as to make recognition not be dependent on transcription by having any subsequent transcription be a separate independent operation apart from of recognition.

23.   Regarding claim 13, the claim is set forth with the same limits as claim 12.

Gupta *et al* (column 7 lines 29-33) reads on the feature of *generating a plurality of*

possible phonetic descriptions, using at least one model (column 4 lines 19-21) *to score*

each possible phonetic description (column 5 lines 3-8) *and selecting the possible*

phonetic description with the highest score as the speech-based phonetic description

(column 5 line 16-18).

24.   Regarding claim 14, the claim is set forth with the same limits as claim 13.

Gupta *et al* (column 9 lines 47-62) reads on the feature of *using an acoustic model* (of

allophones, column 9 line 59) *and a language model* (using *linguistic rules*, column 9

line 38).

25.   Regarding claim 15, the claim is set forth with the same limits as claim 14.

Gupta *et al* reads on the feature of *using a language model comprises using a language*

model that is based on syllable-like units (with the *sub-word units* of column 9 line 62).

26.   Regarding claim 16, the claim is set forth with the same limits as claim 15.

Gupta *et al* (column 10 lines 6-7) reads on the feature of *generating acoustic model*

scores for each of the phonemes in a syllable-like unit & (in column 10 lines 15-18)

summing the acoustic model scores of the phonemes to generate an acoustic model

score for the syllable-like unit.

27.   Regarding claim 17, the claim is set forth with the same limits as claim 12.

- Gupta *et al* (column 10 lines 64-66) reads on the feature of *generating a score for the text-based phonetic description based on the correspondence* (column 11 lines 29-31) *between the text based phonetic description and the representation of the speech signal;*

- Gupta *et al* (column 12 lines 10-16) reads on the feature of *generating a score for the speech-based phonetic description based on the correspondence between the speech-based phonetic description and the representation of the speech signal*

- Gupta *et al* (column 14 lines 24-27) reads on the feature of *selecting the phonetic description with the highest score.*

Gupta *et al*, Häb-Umbach *et al* & Contolini *et al*

28.   Claim 18 is rejected under 35 U.S.C. 103(a) as being unpatentable over Gupta *et al* in view of Häb-Umbach *et al* and further in view of Contolini *et al*.

29.   Regarding claim 18, neither Gupta *et al* nor Häb-Umbach *et al* specify a product reaching a state accessible for human intervention, so do not produce audible pronunciations.

- Contolini *et al* (by selecting the *speaker icon* at the left of figure 2) reads on the feature of *receiving an instruction to generate a audible pronunciation of a phonetic description previously added to the speech recognition lexicon,*

- Contolini *et al* (column 4 line 52-56) reads on the feature of *retrieving the added phonetic description from the speech recognition lexicon, causing an audible pronunciation to be generated based on the retrieved phonetic description* (column 4 line 40).

- It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Contolini *et al* to the device/method of Gupta *et al* & Häb-Umbach *et al* so as to evaluate generated speech.

Schultze & Gupta *et al*

30.   Claims 19-21 are rejected under 35 U.S.C. 103(a) as being anticipated by Schultze (U.S. Patent 6,167,369 A) in view of Gupta *et al*.

31.   Regarding claim 19, the features of the *automatic language identification using both N-gram and word information* of Schultze reads on the *speech recognition system having a language model generated through a process* of the immediate application as follows:

- Schultze does teach the feature of *breaking each word in a dictionary into units* (column 2 lines 55-60) for *n-grams* but does not specify that the units be strictly *syllable-like*. Gupta *et al* reads on the feature of *breaking each word in a dictionary into syllable-like unit* (with the *sub-word units* of column 9 line 62, with the stipulation given in column 3 lines 53-67). Schultze (column 1 line 29) then teaches the further

feature of *for each word* and teaches (by *retaining trigrams* – column 1 lines 24-34 –

into *sequences*) *grouping the syllable-like units of the word into n-grams.* It would

have been obvious to a person of ordinary skill in the art of speech signal processing

at the time of the invention to apply the method/teachings of <u>Gupta *et al*</u> to the

device/method of <u>Schultze</u> so as to separate the contiguous signal into discrete

portions corresponding to the dictionary for match processing.

- <u>Schultze</u> (column 12 lines 21-22) reads on the feature of *counting the total number*

  of n-gram occurrences in the dictionary;

- <u>Schultze</u> (column 12 lines 40-41) reads on the feature of *for each n-gram, counting*

  the number of occurrences of the n-gram in the dictionary and dividing this count by

  the total number of n-gram occurrences to form a language model probability for the

  n-gram.

32.   Regarding claim 20, the claim is set forth with the same limits as claim 19.

<u>Schultze</u> (column 12 lines 35-37) reads on the feature of *breaking the words by*

preferring syllable like units that occur more frequently in the dictionary over syllable-like

units that occur less frequently.

33.   Regarding claim 21, the claim is set forth with the same limits as claim 20.

<u>Schultze</u> (column 12 line 40) reads on the feature of *updating the frequencies of the*

syllable-like units into which the word is broken.

Conclusion

34.   The prior art made of record and not relied upon is considered pertinent to
applicant's disclosure.

-   Sabin *et al* (Japan Patent 11-352994 A) statistical sequence model generator,
    statistical language model generator, and speech recognition system.

-   Kanevsky *et al* (U.S. Patent 6,009,392 A) training speech recognition by matching
    audio segment frequency of occurrence with frequency of words and letter
    combinations in a corpus builds n-gram language models (column 11 lines 10-23).

-   IBM® (IBM Technical Disclosure Bulletin NN87013641, January 1987) procedure for
    using contextual information to obtain improved estimates of word probabilities in a
    speech recognition system introduces building n-gram language models.

35.   Any inquiry concerning this communication or earlier communications from the
Examiner should be directed to Daniel A. Nolan at telephone (703) 305-1368 whose
normal business hours are Mon, Tue, Thu & Fri, from 7 AM to 5 PM.

If attempts to contact the examiner by telephone are unsuccessful, supervisor
Richemond Dorvil can be reached at (703)305-9645.

The fax phone number for Technology Center 2600 is (703)872-9314. Label
informal and draft communications as "DRAFT" or "PROPOSED", & designate formal
communications as "EXPEDITED PROCEDURE". Formal response to this action may
be faxed according to the above instructions,

or mailed to: P.O. Box 1450
                       Alexandria, VA 22313-1450

or hand-deliver to: Crystal Park 2,
                       2121 Crystal Drive, Arlington, VA,
                       Sixth Floor (Receptionist).

Any inquiry of a general nature or relating to the status of this application or

proceeding should be directed to Technology Center 2600 Customer Service Office at

telephone number (703) 306-0377.

                                            Daniel A. Nolan
                                            Examiner
                                            Art Unit 2654
DAN/d
December 4, 2003

                                            DANIEL NOLAN
                                            PATENT EXAMINER